1945

A comparison of scores of deaf and hearing children on the Hiskey Test of learning ability and on performance scales

Jane Gaertner MacPherson

WASHINGTON UNIVERSITY

Central Institute for the Deaf

A COMPARISON OF SCORES OF DEAF AND HEARING CHILDREN ON THE HISKEY TEST OF LEARNING ABILITY AND ON PERFORMANCE SCALES

BY

Jane Gaertner MacPherson

A dissertation presented to the Board
of Graduate Studies of Washington
University in partial fulfilment
of the requirements for the
degree of Master of Science
in Education

June, 1948

Saint Louis, Missouri

TABLE OF CONTENTS

CHAPTERS

APPENDIX

TABLES

FIGURES

Chapter I

INTRODUCTION

In 1941 a new type of test was devised by Hiskey to measure the learning ability of deaf and hard-of-hearing children. In the opinion of Hiskey, it is essential that teachers of the deaf and administrators of schools for the deaf have a test standardized on deaf children alone--a test that will give a better understanding of the ability of young deaf children than can be obtained from intelligence tests standardized on hearing children.

However, it is of vital importance to all who educate the deaf orally, that test scores of deaf children be compared with those of the hearing. The ultimate aim of those using oral methods, is the preparation of deaf children to fit adequately into a hearing environment educationally, and socially as soon as possible. The present investigation, therefore, has been made to study the Hiskey test in its relationship to standardized performance tests measuring intelligence and to apply the Hiskey test to children with normal hearing.

Hiskey termed his test a test of Learning Ability since numerous items have been adopted because of their similarity to the abilities the deaf child must exhibit in school. For this reason, he used Learning Age and Learning Quotient instead of Mental Age and Intelligence Quotient.

Hiskey has made no attempt to compare the intellectual development of deaf and hard-of-hearing children with that of hearing children since he felt that many studies have been made on such comparisons with a lack of agreement and that a study made with his scale could add little if anything to the conclusions reached through other studies. He believed that the important question was not how the deaf child ranked in comparison with the hearing child but how the deaf child ranked in comparison with other deaf children of his chronological age. According to Hiskey, the deaf child's training will never be identical with that of the hearing child, and therefore, the major problem is to find out as much as possible about this child and to organize his educational training so that he will be a better adjusted individual as a result of a more efficient program.[1]

It is true that the initial steps in the education of the deaf are of necessity slightly different from that of the hearing; but after the development of speech and language, it is important to know what the deaf child's chances for success will be when he is placed in competition with hearing companions.

The first tests made by psychologists were not intended as measurements of intelligence, but could more correctly be termed isolated individual tests. They seemed to

---

[1]Hiskey, Marshall S., A Non-Verbal Test of Learning Aptitude Especially Designed for Young Deaf Children, an Unpublished Dissertation, University of Nebraska, Lincoln, Nebraska, July, 1940, p. 8.

have arisen as a result of the individual differences noted by the experimental psychologist in the laboratory. These individual differences were a hindrance to the psychologist at first. Soon he became interested in them for their own sake, and thus the test had its origin. The first tests were concerned with the measurement of specific capacities or abilities and were tests of different mental processes. The mental scale grew out of the individual tests, and is merely a grouping together of individual tests in order to give a more general profile of the mental make-up of the individual.[2] Testing has grown and has gained in momentum up to the present day work on refinement of procedures, careful standardization, and detailed statistical analysis. Much work has been done on the use of intelligence tests in a wide variety of studies attempting to discover the relationship of intelligence to every conceivable human function.[3] +

Many have attempted to define intelligence. Boynton gave the essentials which any definition of intelligence should include:

> First, it must make of intelligence a capacity, and one must realize that, as a capacity, intelligence is part of the native endowment, is unaltered by environment, is not a motivating force or influence, and from the purely biological point of view is probably little more than

---

[2]Pintner, R. and Paterson, D., A Scale of Performance Tests, New York: D. Appleton and Co., 1931, p. 1-25.

[3]Symonds, Percival, "Review and Preview", Rev. Educational Research, 8:217-220, 1938.

> capacity for neural response to organic needs. Second, intelligence cannot be adequately defined in terms of adjustment or adaptation alone, but must incorporate the idea of reconstruction or reorganization. It must recognize the fact that environment is changed by the intelligent individual just as truly as is this individual changed by environment. Third, the best available criterion which we have of intelligence is the behavior of the individual in his group, the fulfilling of group needs. Fourth, it must be seen that intelligence rises with ability to foresee the ultimate needs of the group, to look beyond the temporary to the complete and larger needs.[4]

Binet's definition emphasized three phases of behavior:

> (1) the ability to take and maintain a given mental set; (2) the capacity to make adaptations for the purpose of attaining a desired end; and (3) the power of auto-criticism.[5]

Stern defined intelligence as:

> A general capacity of an individual consciously to adjust his thinking to new requirements.[6]

Thorndike maintained that three distinct types of intelligence must be recognized: mechanical intelligence, social intelligence, and abstract intelligence.

> By "mechanical" intelligence he means the ability to learn to understand and manage things and mechanisms. By "social intelligence he means the ability to act wisely in human relations, and by "abstract" the ability

---

[4]Boynton, Paul L., Intelligence Its Manifestations and Measurement, New York: D. Appleton and Co., 1933, p. 19.

[5]Pintner, R., Intelligence Testing, New York: Henry Holt and Co., 1923, p. 54.

[6]Ibid., p. 54-55.

to understand and manage ideas and symbols.[7]

Colvin believed:

An individual possesses intelligence in so far as he has learned, or can learn, to adjust himself to his environment.[8]

Pintner said:

We need only add a useful way of thinking of intelligence, particularly for the teacher, namely, as ability to learn. For the intelligence test is a good index of a child's ability to learn in school, if he is interested and willing.[9]

Terman said:

An individual is intelligent in proportion as he is able to carry on abstract thinking.[10]

Schick discussed a definition of intelligence as follows:

If we define intelligence as the ability to think abstractly, as the ability to manage ideas and symbols, we shall probably never find an adequate measure for the deaf child because any test of this type involves linguistic ability. Furthermore all school grades and teachers' estimates are weighted on the side of "abstract" intelligence. If, however, we define intelligence as the ability to use judgment and adjust to various situations presented by the environment, then a performance test should give us a measure of the kind of intelligence that is important in our practical work, and a measure not so affected by school training as are linguistic tests.[11]

---

[7]Drever, James, and Collins, Mary, Performance Tests of Intelligence, Edinburgh: Oliver and Boyd, 1928, p. 9.

[8]Pintner, op. cit., p. 56.

[9]Ibid., p. 57.

[10]Ibid., p. 56.

[11]Schick, Helen, "A Performance Test for Deaf Children of School Age", Volta Review, 36:657-659, 1934.

There are two schools of thought concerning theories of intelligence. Some psychologists have held the view of Spearman's theory, that all intellectual activities have in common an important factor which is the essential element of intelligence. This factor is supplemented by many specialized abilities of a narrower range. A person's success in any intellectual performance is the joint product of two factors: specific ability for the task in question and the general ability of the person. Other psychologists have been convinced that no central ability exists, and that

> intellectual adjustments are mediated by a large number of elementary abilities.[12]

Thurstone formulated a new conception of intelligence. He believed it was of practical and scientific importance to isolate those elements of intelligence which were in some fundamental sense primary. Multiple factor methods were developed for the solution of this problem.

> The multiple factor analysis of mental endowment starts with the assumption that if several tasks require the same primary abilities for an effective performance, then the abilities of an individual will not be differentiated by these tasks. On the other hand, if several tasks require different fundamental abilities, it should be possible to differentiate people's abilities by performances on different tasks.[13]

---

[12] Thurstone, L.L., "A New Conception of Intelligence", Ed. Record 17:442, 1936.

[13] Ibid., p. 446.

> As a first approximation it is assumed in factorial analysis that a person's objective performance in a test can be regarded as a sum of the contributions of his several abilities. Some of these abilities may be rather heavily weighted in a particular test, while others may have only slight weight or be entirely absent.[14]

Through factor analysis Thurstone isolated seven primary mental abilities: number facility, word fluency, visualizing, memory factor, perceptual speed, induction, and verbal reasoning.[15]

Rockwell believed that we do not test intelligence directly. Instead

> we measure it through the testing of short samples of learned behavior.[16]

These samples must have been thoroughly learned by a considerable portion of the age level in question. Test standardization demanded that. In test standardization, tests that showed no age discrimination, and those which could not be passed by fifty percent of the children of that age level were rejected.

> Might not this elimination of all content except that which is familiar to at least fifty percent of the children result in an emphasis on overlearned responses?[17]

---

[14]Ibid., p. 445.

[15]Ibid., p. 441-450

[16]Rockwell, John G., "Intelligence Testing: Its Basic Assumptions and Unanswered Questions", Educational Method, 19:81, November, 1939.

[17]Ibid., p. 84.

Rockwell believed that it was possible that one reason intelligence tests show high reliability was the use of very stable overlearned material. He also argued that mental testers use the pass-fail criterion, and this cannot reflect the fine gradations of performance. As a consequence, variability, even if it were present, would be difficult to detect.[18]

Rockwell raised additional question:

> Is it possible that this high degree of reliability results from the fact that in our test-retest procedure we tend to sample within short periods of time rather than at long intervals, and as a consequence may tend to rule out any extreme variation for the mass of subjects?[19]

Most statements concerning the Intelligence Quotient constancy stem from a paragraph of Burks' in 1928 which said:

> Changes of I.Q. of as much as twenty points probably occur only once or twice in a thousand times.[20]

The most extensive of reports on the constancy of the Intelligence Quotient was made by Thorndike. Thorndike, at variance with Burks' results, gave retests after an interval of two and one-half years for over eleven hundred children in three New York private schools, and found the Intelligence

---

[18] Ibid., p. 80-92.

[19] Ibid., p. 89.

[20] Stoddard, George, "The Intellectual Development of the Child", Ed. Digest, 5:18, April, 1940.

Quotient increased more than twenty points in sixteen percent of the cases. These were selected children with an average Intelligence Quotient of 118.[21]

Terman retested 915 pupils over a period of from one day to seven years between test one and two. The average of the second tests was 1.7 Intelligence Quotient points higher than on test one. The increase was independent of the level of intelligence, of the time interval which relapsed, and of the age of the subject tested. Nettels retested one hundred pupils after three years, and found the median of the second tests to be 2.5 Intelligence Quotient points lower than on the first test. These figures indicated group tendencies; individual variation was greater.[22]

Tests contain items about which pupils are uncertain and they make guesses. On repetition of the test they may respond to items differently causing a difference in test-retest scores.

> This regression effect, or tendency for measurements of I.Q.'s to regress toward the true measure, as well as actual changes in pupils between tests given some time apart, the selection of test items, the physical and emotional conditions of the student when taking the test, and other factors, introduce errors which prevent an obtained I.Q. from being entirely reliable.[23]

---

[21] Ibid., p. 17-20.

[22] Tiegs, Ernest, Tests and Measurements in the Improvement of Learning, Houghton Mifflin Co., 1939, p. 29-45.

[23] Ibid., p. 33.

> Intelligence cannot be measured directly. Instead, the items of intelligence tests are largely samples of achievement. The significance of the test score rests largely on the assumption that all students have had approximately the same opportunity to learn, and differences in intelligence are inferred from differences in success on the test. Inasmuch as this assumption is not absolutely true, it constitutes another source of error in determining I.Q.'s.[24]

McNemar stated that the method of finding the Intelligence Quotient

> carried the mathematical necessity that high I.Q.'s will vary more than low ones since the standard error of estimate of the standing of an individual child is a function of the variability of the chronological age group corresponding to his mental age and not to that of his own chronological age.[25]

There has been the question of whether or not the Intelligence Quotient should be used. Some teachers believe that it should not because of errors of measurement in getting the Intelligence Quotient, or because different tests may have given different Intelligence Qutients. Others feel that quotients from intelligence tests are better than estimates and judgments of ability. Then too, the Intelligence Quotient has been misused in attempting to make it primarily an instrument of individual guidance.[26]

---

[24] Ibid., p. 34.

[25] Goodenough, F. L., "Can We Influence Mental Growth: A Critique of Recent Experiments", Ed. Record 21:126, 1940.

[26] Tiegs, op. cit., p. 29-45.

Though there are many different kinds of intelligence tests, the non-verbal test of learning aptitude devised by Hiskey seems to be the only one of its kind. Hiskey said:

> There has long been a need for a measuring device which could be used for guidance purposes and which would give the teacher a better understanding of the young deaf child at the beginning of his educational career. Such a test, to be both functional and practical, should not only discriminate at the lower ages but should be based on the types of tasks which the child is actually required to do in school. In other words, it should give a valid indication of the deaf child's learning level.[27]

Administrators of schools for the deaf

> have desired a useful tool which would enable them to compare the child, of whom they were making a special study, with a large group of deaf or hard-of-hearing children of the same chronological age.[28]

Mental tests have been used for the deaf and hard-of-hearing, but, according to Hiskey, the degree of help which such tests rendered the educator depended upon the number and typicalness of the children on whom they were standardized, and also upon the reliability and amount of information about the children which such tests gave. Most non-verbal or performance tests are weak in these respects when used with deaf or hard-of-hearing children.[29] It is the opinion in the

---

[27] Hiskey, Marshall S., op. cit. p. 1.

[28] Ibid., p. 1-2.

[29] Ibid., p. 2.

present investigation that the experienced examiner can gain additional information concerning the deaf subject on a performance test by observing his method of attack of a test and his persistence in following a problem through to completion. The individual tests which make up the Advanced Performance Scale used at Central Institute were standardized on large numbers of cases and a high reliability was found.

The preceding discussion indicates that there is still a lack of agreement on the definition and theory of intelligence, and the use of intelligence tests. Contradictory results have been cited on the constancy of the Intelligence Quotient in test-retest studies, and some investigators have questioned the use of the Intelligence Quotient.

It is Hiskey's opinion that his Learning Test is a different type of test and that it gives an indication of the deaf child's learning level instead of his mental capacity. In the early use of the Learning Test it was apparent that the scores from the Learning Test and a Performance test were quite similar and it was believed that perhaps Hiskey's Learning Test was really another kind of Intelligence Test.

With the possible introduction of the Hiskey test to many schools for the deaf, the present investigation has been made to obtain further information concerning the reliability and validity of the test; to compare hearing and deaf children on this scale; and to analyse the test items in terms of increasing difficulty as related to increased chronological age and comparative order of difficulty of test items for hearing and deaf subjects.

Chapter II

HISTORY

Mental measurement and intelligence tests developed through four stages according to Morris.[1] The first was a long period of invention when workers were devoted to devising tests and test items. Most were measures of simpler mental processes or abilities and included tests of sensory acuity, mental imagery, reaction time, free and controlled association, and memory problems. The period was one in which the single item was used to measure a single mental ability. The second stage was one of evaluation. Comparisons of relationships between test scores and some criterion of mental ability as school marks, or estimates of intelligence were made. The third stage was ushered in by Binet's theory of intelligence measurement by combining a relatively large number of simple tests into a scale which would give a total intelligence score. He considered intelligence to be a complex organization and proposed measuring it by sampling a wide variety of responses. This period also included the development of the many verbal group tests and revisions of the Binet Scale. The fourth stage was exemplified by experimental and statistical studies to determine something about the nature of intelligence and what tests of intelligence were measuring. During this latter period Thurstone did his work on the nature

---

[1] Morris, Charles M., "A Critical Analysis of Certain Performance Tests", Journal of Genetic Psychology, 54:85-105, 1939.

of mental ability. Performance on fifty-six tests could be accounted for on the basis of seven mental abilities or characteristics of intelligence: number facility, word fluency, visualisation, memory, perceptual speed, induction, and verbal reasoning.

Cattell was the first definitely to use the expression "mental test" in 1890. He analysed the problem of psychological measurement. The present-day psychologist is faced with the same problems of attaining certainty and exactness in psychological experimentation, testing of adequate numbers, constancy, inter-dependence and variation of mental processes, application of psychological measurements to the individual, need of standardization, and effect of time and locality on measurements.[2]

The earliest reference to tests of the deaf is in an article by Greenberger.[3] In 1889 he stumbled on the idea of mental tests when faced with the problem of weeding out the feeble-minded among the children applying for admission into an institution for the deaf. He suggested showing attractive books to the child and watching his reactions. If the child remained apathetic it was a bad sign, but if he was attracted by them and maintained an interest in them, it was an indication of a fair mentality. He also proposed a

---

[2] Boynton, Paul L., Intelligence: Its Manifestations and Measurement, New York: D. Appleton and Co., 1933, pp. 152.

[3] Pintner, Rudolf and Peterson, Donald G., "Psychological Tests of Deaf Children", Volta Review, 19:661-667, 1917.

number test, on the assumption that lack of an idea of numbers is a sign of weakness of a child's mind. Other tests consisted of showing forms and colors, and building with blocks. Modern mental tests make use of all these materials. There was no standard method of procedure or evaluation of the child's responses in terms of responses of other children. The essential elements of standardized tests as we know them today were lacking.

Taylor reported on a spelling test given to 148 deaf children and a group of hearing children. It was a free association test in which the children were told to write as many words as they could in fifteen minutes. The average number of words written by the deaf was 161; by the hearing, 153. The average percent of spelling mistakes by the deaf was 2.7, by the hearing 4.3. He concluded that the deaf were superior in spelling ability.[4]

Mott[5] in 1899 compared eight year old deaf and hearing children. She used physical tests and memory tests. In physical measurements, athletic tests, and manual dexterity, the deaf were as good as, or better than, the hearing. Mott referred to mental factors saying:

---

[4]Ibid., pp. 661-667.

[5]Pintner, R., Eisenson, Jon, Stanton, Mildred, The Psychology of the Physically Handicapped, New York: F. S. Crofts and Co., 1941, pp. 101-130.

> Memory and observation seemed the only mental powers capable of exact comparison between the two classes of children.[6]

On these mental factors the deaf were markedly superior to the hearing. This was the beginning of the use of tests comparing the deaf and hearing.

Smith in 1903 tested the memory of the deaf child. Thirty-seven pupils memorized a story in twenty-five minutes, and then Smith tested immediate and delayed recall. The grading was subjective. He concluded that the best pupils, as a rule, had the best memories, and the poor pupils had poor memories. He found very little difference between the immediate and the delayed recall tests in the case of the best pupils; no marked difference between the sexes in ability to memorize.[7]

The study of MacMillan and Bruner is the first one in which standard psychological tests were given to deaf children. They gave the single separate abilities tests which preceded the general intelligence tests of the Binet type. The deaf, on the average, were below the hearing in lung capacity, but there was not much difference between them in strength of grip. In the tapping test twenty-four percent of the deaf reached or exceeded the average of the hearing, while on the cancellation of A's test the deaf on the average were two to three years less mature than the hearing. The

---

[6] Ibid., p. 107.

[7] Pintner, and Paterson, op. cit., pp. 661-667.

deaf were below the hearing on perception of size by sense of touch and equal to them on perception of differences in lifted weights, but were definitely below the hearing on visual memory span for numbers. the authors concluded:

> This inferiority of the deaf on the mental side perhaps means no more than that the child is from three to four years less mature than the hearing child of his own age, and that this date of maturing will be correspondingly delayed. It does not necessarily mean that adult deaf individuals will be much inferior in mental comprehension and initiative to hearing adults. Indeed this mental retardation may be due to conditions in training, and were the deaf child's instruction begun in infancy instead of at six years, the difference might be reduced or even eliminated.[8]

Pintner and Paterson in 1916 used the Goddard Revision of the Binet-Simon Scale. They used any method or combination of methods as speech, writing, signs and manual spelling to explain the directions to the subject. The tests were not satisfactory for use with the deaf. Of the twenty-two children tested only one reached the norm for his age. The average mental retardation was 4.58 years. The language handicap and the lack of experience with certain things handicapped the deaf child on the test.[9]

Since the Binet test proved impracticable, Pintner and Paterson constructed their own scale. The purpose was to

> arrive at an estimate of the pupil's native ability, rather than the knowledge that he has acquired from his school or home environment.[10]

---

[8] Pintner, Eisenson, Stanton, op. cit., p. 109.

[9] Ibid., p. 101-130.

[10] Pintner and Paterson, op. cit. p. 665.

> A complete separation of native ability and acquired knowledge is doubtless impossible, but at least we can strive to approximate the ideal of measuring native ability alone. For this reason psychological tests present relatively new situations to the child to avoid as far as possible anything bearing upon school work.[11]

Pintner and Paterson stated further:

> General Intelligence has been defined as the ability of the organism to adjust itself adequately to relatively new situations. The tests of the psychologist are the relatively new situations to which the child is called upon to adjust itself, and the adequacy with which it can so adjust is the measure of its intelligence.[12]

Some of the first work done by Pintner and Paterson was in 1914 and 1915. Eighteen deaf children age seven from the Ohio State School for the Deaf and fourteen hearing children age six were tested on the Seguin form board on entering school in October 1914 and then retested October 1915. The hearing group were pupils in the first grade of a public school and were of parents of the laboring class. For three trials the average number of errors and the average time were the basis for scoring. From the results the authors concluded that the average entering class of deaf children were apt to be about a year retarded in form board ability and this backwardness was not made up during the first year in school.[13]

---

[11] Ibid., p. 666.

[12] Ibid., p. 666.

[13] Pintner, R. and Paterson, D., "The Form Board Ability of Young Deaf and Hearing Children", American Annals of the Deaf, 16:184-189, March 1916.

About the same time, Pintner and Paterson tested individually four hundred and eighty-one deaf between the ages of seven and twenty-six at the Ohio State School for the Deaf during the school year 1914-1915. The test given was one for visual memory for digits. They compared their results with averages of deaf tested by MacMillan and Bruner in the Chicago Day-School for the Deaf, and also with Smedley's hearing children. The authors concluded that:

1. Oral pupils are superior to manual pupils on the average.
2. Deaf children as a group have an abnormally poor memory span due to the lack of auditory experience.
3. Adventitiously deaf children are superior to congenitally deaf children on the average.
4. Auditory experience plays an important part in the efficiency of both hearing and deaf individuals in visual memory for digits.
5. Sex differences do not exist among the deaf in this test.[14]

After standardization of the scale, and the rejection of certain tests which did not give satisfactory norms, Pintner and Paterson published their battery of fifteen tests in 1917. This was revised and shortened to nine tests by Pintner and Spaid in 1918.[15] Tests were selected to get as wide a variety as possible so that all the various factors entering into intelligence might be brought into play. Each

---

[14] Pintner, R. and Paterson, D., "A Comparison of Deaf and Hearing Children in Visual Memory for Digits", Journal of Experimental Psychology, 2:76-88, 1917.

[15] White, J. L., "A Performance Test for School Age Children With a Language Handicap", Unpublished M.S. Thesis (Washington University, St. Louis, Mo.) 1939, pp. 9-21.

test was to present a relatively new situation to the child, and no verbal instructions were to be necessary in order to give the tests. Signaling to go ahead quickly was all that was necessary.[16]

Results on the scale indicated mental inferiority of the deaf of from two to three years. There was no difference between the congenitally and the adventitiously deaf. Pintner believed that the mental retardation was due to the same cause that produced the deafness.[17]

Hiskey criticized the Pintner and Paterson test. He said it was standardized on hearing children and this made the use of the norms for the deaf a questionable procedure. It was heavily weighted with form boards. Since many tests are timed, and scored in terms of speed of performance it placed the deaf at a disadvantage because of the difficulty of indicating the need for speed to the young deaf.[18]

Several other investigators developed performance tests of various types, but the lack of adequate standardization on most of them has prevented an understanding of the results obtained thus far. Healy and Fernald devised a group of tests, some of which have been used by later workers. They were not grouped in a scale nor were they standardized.

---

[16]Pintner, R. and Paterson, D., "A Scale of Performance Tests, New York: D. Appleton and Co., 1931, pp. 1-25.

[17]White, op. cit. pp. 9-21.

[18]Hiskey, M. S., A Non-Verbal Test of Learning Aptitude Especially Adapted for Young Deaf Children, A Dissertation, Lincoln, Nebraska, July 1940, p. 2.

//Knox constructed a rough scale of performance tests for immigrants at Ellis Island, but they too were not standardized. However, some of the items were used extensively by others and norms established on them.[19]

Kohs designed his Block Test in 1923. He advocated the incorporation of his test in a series of performance tests which might be made to serve as a check on Binet results.[20]

Stutsman developed a series of performance tests at the Merrill-Palmer School in Detroit for testing children ranging from seventeen to sixty-six months of age. She used older tests as the Wallen Peg and the Goddard Form Board together with a group of new ones.[21]

Bayley's eight tests were for 3, 4, and 5 year olds. Four of her tests--auditory memory, visual memory, picture perception, and naming objects--are less manipulative than are most performance tests. There was a correlation of .90 with the Binet.[22]

Gaw studied intercorrelations of performance tests. The fact that the intercorrelations among the performance tests were rather low made it very necessary to use a number of tests to arrive at a mental age, just as the Binet scale

---

[19] White, op. cit. p. 9-21.

[20] Boynton, op. cit. p. 149-206.

[21] Ibid., p. 149-206.

[22] Ibid., p. 149-206.

combined several tests. The best tests of this central performance capacity were: Cube construction, Substitution, Goddard Form Board, Picture Completion Tests I and II, and possibly the Profile Test. The least effective were the Dearborn Form Board, the Manikin, the Adaption Board, Healy A, and the Triangle Test.[23]

The Pintner Non-language mental test was used in two large surveys: (1) the Reamer survey in 1921 of twenty-six schools, using 2172 children from eight to twenty-one years old and over; (2) the National Research Council survey in 1928 of forty-one schools using 4432 pupils twelve to twenty-one years old and older. The latter study was made by Day, Fusfeld and Pintner. The average of the four Intelligence Quotients from the National Research Council survey (ages twelve to fifteen) was 85.8; the average of the five Intelligence Quotients from the Reamer survey (ages twelve to sixteen) was 88.4. An Intelligence Quotient of eighty-six might be considered as an approximate estimate of the intelligence of deaf children on group non-language intelligence tests. Both studies found the age of onset of deafness had an influence on the score of the test. Both studies found the intelligence of manual students to be the lowest; Reamer finding the combined method students to be highest and the Council survey finding oral pupils highest.[24]

---

[23] Ibid., p. 149-206.

[24] Pintner, Eisenson, Stanton, op. cit. p. 101-130.

In the period following that of the early investigators many studies were made using the performance test. Boynton said that the performance test did not correlate as highly with scholastic achievement as did the Stanford and Herring revisions of the Binet, for example.

> Here, though, it may well be asked as to whether it is desirable for the performance tests to correlate so highly. No one knows just exactly the true correlation between scholastic achievement and native ability and the fact that the performance test does not show a closer positive relationship either with scholarship or with the results of the other tests, in truth, may be its more desirable feature.[25]

In 1928 Drever and Collins administered their scale to four hundred hearing children and deaf children. This scale was an organization of tests designed by other investigators and included: Kohs Block, Knox Cube, Domino Test, Size and Weight Test, Manikin and Profile, Two Figure Board and Healy's Puzzle A, Cube Construction, Healy's Picture No. 1, and Little Bo-Peep Picture Completion. The general norms for hearing children (ages six to sixteen) were given, and for purposes of comparison the norms based on deaf subjects alone (ages six to sixteen) were also given. Both sets of norms were to be regarded as provisional. From the results of their tests Drever and Collins concluded:

> In no single case are the deaf as much as one year retarded. These results, therefore, so far from confirming the general view that the

---

[25]Boynton, op. cit. p. 195-196.

> deaf child is about three years retarded as compared with the normal, show rather that, when the language factor is entirely eliminated from our tests, it is very doubtful whether the deaf child is retarded at all. At any rate, it appears certain that the deaf child is much less retarded than has been generally supposed.[26]

Hiskey referred to the use of the Drever and Collins test on fifteen hundred hearing and deaf children with norms based on the results of the combined group. This later work had been done since the original testing in 1928.[27]

In 1930 Arthur published a point scale of performance tests standardized on eleven hundred public school children from middle-class American homes. All the items were from other well known tests, but the author stipulated that the order of presenting the items must be rigidly observed, since this factor would affect the norms. The chief difficulty in constructing the scale was to find items sufficiently difficult to test the superior adult without involving school training or experience, and she recommended further studies on this problem.[28]

Peterson and Williams gave the Goodenough Intelligence test to three hundred thirty children in five schools for the deaf varying in methods of instruction. They ranged

---

[26] Drever, James and Collins, Mary, Performance Tests of Intelligence. A Series of Non-linguistic Tests for Deaf and Normal Children, Edinburgh: Oliver and Boyd, 1928, p. 43-44.

[27] Hiskey, op. cit. p. 3.

[28] White, op. cit. p. 9-21.

in ages from five to fourteen years. The curve of distribution did not follow the normal curve for hearing children, for there were a greater number clustered at the lower portion of the curve. There was a normal decline at the upper portion. If the mean, median, and range of the deaf were advanced twenty points on the scale of Intelligence Quotients, the two curves would almost coincide. This twenty point retardation considered in terms of years finds a retardation of the deaf to range from one year at age five to two years seven months at age thirteen, increasing proportionately according to the advancing age of the children. The teachers' estimate of intelligence yielded low correlation with the Intelligence Quotients obtained.[29]

Schick used the Randall's Island Performance Series in October 1931 and October 1932. In 1931 the median score for thirty deaf children was 95.0, for thirteen children in the speech department the median score was 104.0, and for the total forty-three, it was 96.0. In 1932 the median score for twenty-nine deaf children was 97.0; for thirty children from the speech department and speech clinic, 97.5; and for the total fifty-nine the median score was 97.0. The correlation coefficient was computed between the Intelligence Quotients of the thirty-five children retested and found to

---

[29]Peterson, E. G. and Williams, J. M., "Intelligence of Deaf Children as Measured by Drawings", American Annals of the Deaf, 75:273-290, 1930.

be .91 with a probable error of $\pm.02$, indicating high reliability. Schick called attention to the average Intelligence Quotient of 97.0

> which is proof that the deaf and speech defective children are not mentally retarded, but only retarded educationally because of their language handicap.[30]

Soyster at the Illinois School for the Deaf tested fifty deaf children on the Merrill-Palmer Scale, the Minnesota Preschool test and the Pintner-Paterson scale and found no retardation. There was a close correlation with ability in school subjects.[31]

Goodenough and Shirley administered the Goodenough Drawing test and the Pintner Non-Language test to 406 deaf children in Minnesota between the ages of five and twenty years and over. The median Intelligence Quotient on the former was 87.7 and on the latter 98.4.[32] They found a correlation of .33 between the Goodenough and the Pintner Non-language for 229 cases and concluded they do not have much in common.[33]

The Lectometer, designed by Meyer, gave results correlating highly with a Binet on hearing children. The

---

[30]Schick, Helen, "The Use of a Standardized Performance Test for Pre-School Age Children With a Language Handicap", Proceedings of the International Congress on the Education of the Deaf, West Trenton, N.J., June 1933, p. 532.

[31]Cutler, E. M., "Summary of Psychological Experiments with the Deaf: 1932-1938", American Annals of the Deaf, 86:181-192, March 1941.

[32]Ibid., p. 181-192.

[33]Pintner, Eisenson, Stanton, op. cit. p. 101-130.

task imposed by the instrument requires the subject to copy a series of patterns of dots from a row of blocks onto a row of corresponding tablets made up of tumbler switches. The pattern had to be reproduced in an up-down reversal, and the same principle of reversal of pattern was applied throughout the test.[34] The instrument was used for testing two hundred deaf children (ages five to twenty-one years and over), and results showed a distribution of Intelligence Quotients following the normal curve, with a range from feeblemindedness to genius. The median quotient was around one hundred.[35]

// In 1933 MacKane made an investigation to see if the findings of Drever-Collins were correct since they differed so from other studies. He used the Pintner-Paterson Performance scale, Drever-Collins Performance scale, Arthur Performance scale (making changes to give it without verbal directions), and the group Pintner Non-language mental test. He matched one hundred thirty deaf, ages ten to twelve, with hearing subjects for chronological age, sex, socio-economic status, and nationality of parents. The results supported those of Drever and Collins in general. At no age level were the deaf as much as one year

[34]Schick, Helen F., and Meyer, Max F., "The Use of the Lectometer in the Testing of the Hearing and the Deaf", __American Annals of the Deaf__, 77, 4:292-304, September, 1932.

[35]White, op. cit. p. 9-21.

retarded. He found no superiority of the deaf. The Pintner Non-Language test showed a marked superiority of the hearing which convinced MacKane that it does not measure the same abilities as the performance scale.[36]

Schick tested fifty deaf children on the Drever-Collins Performance Scale and found a median Intelligence Quotient of 105 or 122 according to the method of scoring.[37] Amoss reported on Intelligence Quotients for 288 deaf children, ages five to twenty-two years, on the Ontario School Ability Examination. The range of Intelligence Quotients was from 49-192 with a median of ninety-four.[38] Using one hundred deaf on the Kohs Block Design test Peterson concluded it the most satisfactory single performance test yet designed. The degree of retardation of the deaf was less than on the Goodenough test. That was expected since there was a positive correlation between the Block Design test and the Binet of .84, while that of the Goodenough and the Binet was .76.[39]

The Arthur Performance Scale was given to ninety children in special classes for the deaf and hard of hearing in St. Paul by Bishop. The age range was from five years

---

[36] Ibid., p. 9-21.

[37] Cutler, op. cit. p. 181-192.

[38] Pintner, Eisenson, Stanton, op. cit. p. 101-130.

[39] White, op. cit. p. 9-21.

1936

seven months to sixteen years eleven months. The Intelligence Quotients ranged from 68-152 with a mean of 97.16 for the group. She concluded that her findings indicated a normal distribution with as nearly an unselected group of deaf children as can be found in the general school population.[40]

Springer used three hundred thirty matched pairs of deaf and hearing between the ages of six and twelve years on the Goodenough Drawing test and found a mean Intelligence Quotient of 96.24, indicating very little retardation.[41] On eighty-nine deaf children on the Arthur Performance Scale Kirk found a median Intelligence Quotient of 99.4.[42] Streng and Kirk, using ninety-seven deaf children between the ages of six and eighteen years on the Arthur Performance Scale and the Chicago Non-Verbal test, found a median Intelligence Quotient of 100.9 on the former and a median Intelligence Quotient of 95.5 on the latter.[43] Bowers, at the West Virginia School for the Deaf found the mean Intelligence Quotient for two hundred cases was 89.2. Of nine performance items given the lowest mean rating, 69.29, was on the Mare and Foal test, while the highest,

---

[40]Bishop, Helen, "Performance Scale Tests Applied to Deaf and Hard of Hearing Children", Volta Review, 38:447, 1936.

[41]Cutler, op. cit. p. 181-192.

[42]Ibid., p. 181-192.

[43]Ibid., p. 181-192.

101.25, was on the Porteus Maze.[44]

Zeckel and Van Der Kolk in 1939 used the Porteus test with one hundred hereditary deaf children and one hundred hearing, ages seven to fourteen years, in Holland. The younger hearing group (ages seven to eleven) had an average Intelligence Quotient of 88.64 while the average for the younger deaf group was 78.65. The older group of hearing (age twelve to fourteen) had an average Intelligence Quotient of 108.11, while the older deaf group averaged 95.77. They said the considerable backwardness of the younger deaf group was due to a great extent to the low Intelligence Quotients of girls in this group. However, the curve for each age of hearing children regularly and without exception rose above that of the deaf. The curve for every age of the deaf child was below the hearing and did not intersect or even catch up in later age periods. The authors indicated:

> There are signs which indicate that the lack of perceptions of an auditory nature has an influence upon the development of intellect. Deafness creates a mental backwardness which impedes also the development of regions of the intellect other than those developed by speech.[45]

In 1939 White used the Advanced Performance Scale

---

[44] White, op. cit. p. 9-21.

[45] Zeckel, A. and Van Der Kolk, J. J., "A Comparative Intelligence Test of Groups of Children Born Deaf and of Good Hearing, by Means of the Porteus Test", <u>American Annals of the Deaf</u>, 84:122, 1939.

which was also administered in the present study.[46] Her study will be discussed in detail since it is so closely related to the present one. She administered the scale to 239 children who had been given other standardized and reliable tests. Results were compared with scores on these tests and correlation coefficients were computed. One hundred thirty-three deaf and speech defective cases at Central Institute for the Deaf had been given other performance tests: sixty-five were tested on the Randall's Island scale for younger children and sixty-eight on the Lectometer designed by Meyer. The other one hundred six children were hearing children from speech clinic at Central Institute and from the various public schools and one Sunday School. The linguistic tests used for comparison were the Stanford-Binet, Detroit Kindergarten and First Grade test, Haggerty, Henmon-Nelson test, Kuhlmann-Anderson, Otis, and the Pintner-Cunningham.

On comparing the Performance Scale with the other performance tests, White found a validity coefficient for those tested on the lectometer of .78; for those tested on the Randall's Island series .65; and for all it was .71. The correlation coefficients, which are fairly high, showed this scale measured the same type of intelligence that was measured by other standardized performance tests. The

[46]White, op. cit., Unpublished M.S. Thesis, Washington University, St. Louis, Mo., p. 32-47.

probable errors were very low, indicating high statistical reliability for the test.[47]

The validity coefficients for the Performance Scale and the language tests were not as high as on the other performance tests, but they still showed a fair degree of relationship between the two types of tests. The coefficient for all the language tests was .63. The probable errors of correlation between the total language tests and each separate test (except the Kuhlmann-Anderson) with the Performance Scale showed statistical reliability. The author said:

> On a whole, the comparison of the performance scale with other performance tests and with language tests, shows a high degree of reliability. Although the correlation coefficients are not sufficiently high to justify making any claims as to the prognostic value of the scale, they do indicate that the tests are measuring the same type of ability to a fairly great extent. Within certain limits, they could be used to good advantage for purposes of homogeneous grouping, and a fairly accurate estimate of intelligence. Their chief use with the normal hearing individual would be as a check on the linguistic test. Their value in testing the deaf and speech defective would undoubtedly be increased by being given in conjunction with another performance test of a different type.[48]

Forty-seven and two tenths percent of the language test scores and 57.1 percent of the scores on other performance tests varied less than ten points from scores on the

---

[47] Ibid., p. 32-47.

[48] Ibid., p. 36-37.

Performance Scale. This gives additional evidence that the latter is a valid instrument for measurement of intelligence.[49]

A comparison of the scores of all the groups on the Performance Scale showed the hearing to be decidedly above normal with a median Intelligence Quotient of 113.33. The author believed that the selection of cases had an influence on the scores. The median of the deaf, presumably an average group, was 104.07. The speech defective group had a median of 99.00, but inclusion of scores of several spastics and feebleminded children who have speech defects may have altered this group's rating. However, all three groups were within the normal classification. The insufficient number of cases and the lack of matching made a statement regarding the relative intelligence of hearing, deaf and speech defective groups unjustified.[50]

The mean and median Performance Quotients of the deaf were found for each item included in the Performance Scale to see which tests were most difficult and which were completed easily. The deaf ranked above average on the Block Design test, Manikin-Profile, and Picture Completion tests. These items were intended to test synthesis, analysis and visual perception. It was apparent that the deaf were not deficient in these traits. The below normal

[49] Ibid., p. 32-47

[50] Ibid., p. 32-47

score on Knox Cube was interpreted to mean that the deaf were retarded in memory span. The low score on the Seguin Board seemed to indicate the deaf were below average in perceptual speed. The author thought this might be due to lack of stress of speed in their education, to poor motor coordination of spastics or loss of vestibular function. Form board scores showed retardation in visual perception of an abstract nature. This might be due to the fact that the deaf's early training is concrete in nature, or to poor muscular coordination.[51]

Since only a few studies have been made with learning tests for the deaf, there is a dearth of information concerning them. Pintner, Eisenson, and Stanton said:

> Tests of learning ability may be considered as a type of intelligence test.[52] .....The learning tests cannot be translated into I.Q.'s, but the general results corroborate very well the findings on the intelligence tests proper.[53]

Pintner and Peterson worked with the Digit-Symbol and the Symbol-Digit tests before the introduction of the modern type of group intelligence test. They tested nine hundred and ninety-two deaf with the Digit-Symbol test and one thousand forty-nine with the Symbol-Digit test. They converted the deaf median score for each age into the equivalent hearing age according to Pyle's standardizations for

---

[51] Ibid., p. 32-47.

[52] Pintner, Eisenson, Stanton, op. cit. p. 122.

[53] Ibid., p. 127.

hearing children. There were no hearing norms available for children below eight years of age. They found that at no age did the deaf child equal the norm of the hearing. The deaf were two to three years retarded at those ages where best comparisons could be made. The average retardation on the Digit-Symbol test was 3.75 years and on the Symbol-Digit test 2.9 years. They found no sex differences or differences between the congenitally and adventitiously deaf.[54]

Newlee gave both of the above tests to eighty-five deaf children in the day school in Chicago in 1918, whose age range was from six to eighteen. She compared her results with Pyle's norms and also the norms of Pintner and Paterson for deaf children. The Chicago deaf children were found to be equal to the hearing children in this particular learning ability. Acceleration, retardation, and normal ability of these children were distributed according to the normal curve. There was very little difference between the sexes.

> The results of the tests lead to the conclusion that in native ability to build up habitual association quickly--that is, to learn--these Chicago deaf children are equal to hearing children.[55]

In 1940 Hiskey developed a non-verbal test of learning aptitude in Lincoln, Nebraska. He believed that admini-

---

[54] Ibid., p. 101-130.

[55] Newlee, Clara E., "A Report of Learning Tests with Deaf Children", <u>Volta Review</u>, 21:223, 1919.

strators of schools for the deaf had long been looking for a measuring device which would give them a reliable indication of the learning possibilities of the younger deaf or hard-of-hearing child. They wanted a useful tool which would enable them to compare the child with a large group of deaf or hard-of-hearing children of the same chronological age.[56]

It was Hiskey's opinion that other tests were standardised all or in part on hearing children, or on an insufficient number of deaf or hard-of-hearing children; that the tests were not usable at the lower age levels; that they were over-stocked with the form board type of item; and that they placed more emphasis on the time element than is desirable with young deaf. Hiskey's test was standardized on a group of deaf and hard-of-hearing children. He used four hundred sixty-six of the age range three to ten. He eliminated many speed tests. Items in several instances were selected because of the close relationship to tasks actually done in school.[57]

The small number of cases in the lowest age group meant that the norms were less reliable at age four. The children might be brighter since schools are hesitant to accept young deaf unless they seem rather capable. At each age level except one the mean learning age was slightly

---

[56] Hiskey, op. cit., Chapter I.

[57] Ibid., Chapter I.

higher than the mean chronological age. Hiskey felt that this was a desirable feature. If there were an inadequate sampling of subjects it would probably apply to the group with limited ability, since the duller are less likely to have entered school at a reasonably early age than the brighter. The fact that the test ceiling is not high might be responsible for the slightly lower mean learning age for the group at the upper end of the range.[58]

The test of Hiskey seems to be in a class by itself, for in none of the literature is a test of its kind listed. Hiskey himself says it was impossible to find the validity of his test through correlation with other test scores since there was no existing test which would have been an acceptable criterion. He intends to make a further study based on the relationship between the test results and teacher judgments.[59] //

A review of the literature for the past three years showed that "no unexpected data were reported for exceptional groups."[60]

Recent studies have shown less retardation for the deaf than the earlier work indicated, but the difference in results is still too great to justify drawing any significant conclusions.

---

[58]Ibid., Chapter III.

[59]Ibid., Dissertation.

[60]Freeman, Frank S., "Psychological Tests and Their Uses", Review of Educational Research, 14:33, February 1944.

Chapter III

DESCRIPTION OF THE TESTS

Three scales were used in the present investigation: the Advanced Performance Scale for children six years and above; the Randall's Island Performance Series; and the Nebraska Test of Learning Aptitude for Young Deaf Children constructed by Hiskey.

All of the items used in the Advanced Performance Scale have been previously standardised in the form in which they are given here. The tests selected are all included in the well-known performance scales. The sequence of presenting the items is that of Drever and Collins with the Seguin Board added in the order of the Arthur Scale. The method of scoring is in each case the same as that used by the originator of the test.[1]

All directions may be given in pantomime. With hearing children they are given verbally for the sake of naturalness. Pintner and Peterson also said about their scale:

> Naturally in giving the test to hearing children the examiner will say something, but what he says is not essential for the understanding of the test. If the examiner in testing a hearing child were not to say anything, he could introduce an embarrassing and abnormal element into the situation.[2]

---

[1]White, Jennylouise L., "A Performance Test for School Age Children with a Language Handicap", M.S. Thesis, Washington University, 1939, p. 22.

[2]Pintner, R. and Peterson, D., A Scale of Performance Tests, New York: D. Appleton and Co., 1931, p. 22.

The time required to give the Advanced Performance Scale is from thirty-five to seventy minutes with an average of about forty-five minutes.

Test 1--Block Design--(The following description is taken from previous work done on this scale at Central Institute for the Deaf.) The Block Design test was designed by Kohs to test mental analysis and synthesis, and is included in the Drever and Collins, Arthur, and MacKane scales. It consists of sixteen one-inch wooden cubes each painted in the same way, the sides being red, white, blue, yellow, blue and yellow, and red and white. The sides with the two colors are divided diagonally. Seventeen designs drawn in the same colors on small cards are presented in increasing order of difficulty for the subject to construct.

The procedure of Drever and Collins is followed to explain the test. Four of the sixteen blocks are placed in front of the subject. The examiner turns the four blocks over until all four blue sides are up. Then the blocks are turned to show the other sides: red, yellow, white, red and white, and blue and yellow. The examiner turns the first block and the child matches the remaining three to this. This is done to show the child all sides of the blocks, and to make clear to him that all the blocks are the same. The examiner then constructs the first pattern as a sample, scatters the blocks, and indicates that the child should put them together. If he is able to do this within the assigned time limit, the examiner gives him the next pattern. The test is scored on the basis of the number of moves and the time taken for completion of each pattern.[3]

Test 2--Knox Cube--Knox devised this test for memory span and it is used, either in the original form or with slight variations, by Drever and Collins, Pintner and Paterson, Arthur, and MacKane. The material consists of six one-inch wooden cubes painted black. Four blocks are placed in front of the subject in a row, about two inches apart. The other two are used by the examiner and the subject for tapping. Some investigators use only one cube for tapping and pass it back and forth from the examiner to the subject. This procedure causes loss of time and distraction of the subject's attention, which might in turn cause failure on the particular item. The introduction of a sample pattern, 1-2-3, is another slight variation. This is added to be sure the subject understands he is to imitate the pattern tapped by the examiner. This sample pattern is used by Maxfield of Ohio State University as a substitute for the verbal directions given in testing manuals:

Watch carefully, and then do as I do. Do that.

The examiner taps a pattern at the rate of about one tap per second and then indicates to the subject that he is to imitate the pattern. There are twelve patterns used and they are arranged in the order used by Drever and

---

[3] White, op. cit. p. 24.

Collins, which is slightly different from that used by Pintner. Three failures are allowed before the test is discontinued. The test is scored according to the number of patterns successfully completed.[4]

Test 3--Seguin Board--This test, standardized by Sylvester, was used by Pintner and Paterson, Drever and Collins, Arthur and MacKane. It consists of ten blocks of common geometrical shapes, which are to be placed in the appropriate place on the board at a given signal. Three trials are given, and the shortest time is used for scoring. The blocks are arranged before each trial in the manner specified by Sylvester.[5]

Test 4--Manikin and Profile--Both of these tests for synthesis were included in all four of the well-known scales. The Manikin test was devised by Pintner and is used for younger children as he has suggested. The material consists of a figure of a man, cut from wood, and divided into a head, the body, the two arms and two legs. The examiner arranges the pieces in the manner described by Pintner and indicates that the subject is to put them together. He gives no indication of what the finished product will represent. A maximum of five minutes is allowed. The test is scored according to the accuracy of the assembled figure. Perfect performance scores a mental age of seven years, and therefore, is not

---

[4] Ibid., p. 24.

[5] Ibid., p. 25.

useable for older children. If the subject makes a perfect score on this test, it is simply recorded as completed and he is given the Feature Profile.

The Feature Profile was devised by Knox and consists of a human head cut out of wood and divided into eight pieces. Scoring is based on the time required to complete the profile, but there is a time limit of five minutes. The subject is not told what the finished figure is intended to represent, in accordance with Pintner and Paterson's procedure.[6]

Test 5--Form Boards--This section of the test consists of three form boards: the Two-Figure Board, the Healy Puzzle A, and the Casuist. Each test is scored individually and then the mean score of the three is recorded as the score on the whole section. A maximum of five minutes is allowed for each test, but the score is determined on a basis of time required for completion.

The Two-Figure Board, which was devised by Pintner, is used in all of the standardized scales. The board has a square and a cross cut out, and the subject must fit the nine pieces into it. The pieces are arranged above the board.

The Healy Puzzle A, designed by Healy and included by Drever and Collins, MacKane, and Pintner and Paterson in the longer version of their performance series, is of the

---

[6] Ibid., p. 25-26.

specifications given by Pintner and Peterson. The puzzle consists of a frame, into which five rectangular pieces are to be fitted.

The Casuist Board, used in all the scales except that of Drever and Collins, was devised by Knox. The board is rectangular with three circles of varying sizes and an oval. Into these recesses twelve pieces must be fitted.[7]

Test 6--Picture Completion--Healy's Picture Number I is used in each of the four standardized scales. It is a colored picture of an outdoor scene. Ten different activities are represented. A one-inch square is cut out in connection with each of these ten activities. The subject is to select from fifty squares the one most suitable for insertion. Most subjects require no longer than five minutes, but ten minutes may be allowed if required. The score is determined only by the correctness of the insertion. There are several possible solutions for each insert, but there is only one correct one that gives full credit. Selection of one of the other possible choices gives part score on the total score. Mental age is determined by the total number of points, following Healy's norms.[8]

Test 7--Drawing Test--This test, adapted from the Binet Scale, is a supplementary test used as a check with subjects

[7] Ibid., p. 26-27.

[8] Ibid., p. 27.

whose ability has not exceeded a seven year average on the preceding tests. The examiner shows the subject a card with a circle drawn on it and indicates that he is to draw three of them. If one of the three is acceptable he continues in the same way with the square, triangle, inverted triangle, and diamond until he fails one of the tests. An acceptable circle scores three years, a square scores four years, the triangle scores five years, inverted triangle scores six years and the diamond seven years.[9]

After finding the mental age on each of the six tests (or seven if the drawing test is used), the mental age for the whole scale is determined by calculating the mean of these scores.

Randall's Island Performance Series is the outcome of the cooperation of the Department of Psychology at Columbia University and Psychological Laboratory of Children's Hospital and Schools on Randall's Island, a New York City Institution. The test material has been assembled from the performance tests of Stutsman, Pintner and Paterson, Gesell, Baldwin and Stecher and standardized to form a scale which will, according to the authors:

> (a) give as broad a clinical picture as can be obtained without the use of language; (b) give a numerical rating which will correlate highly with a Binet rating; (c) give a median mental age of the same group by a Binet rating; (d) be applicable to older retarded as well as to

---

[9]Ibid., p. 28.

unselected young children; (e) be within the fatigue limit and attention span of children of mental ages two to five years.[10]

The test was constructed on a continuous scale. The tests were grouped according to dominant mental activity. The use of language in presenting the tests is incidental, and does not affect the results.

Schick at Central Institute for the Deaf, made several alterations in standard procedure to clarify the task for the child with no language. All changes in procedure were substitutions for verbal instruction, and she did not feel that the alterations were significant in affecting the norms of the standardized tests.[11] Except for these alterations, standard procedure was followed in the present study.

The series has been divided into nine sections: manual planning, manual dexterity, imitation of movements, form perception, eye-hand coordination, adaptation, color sorting, counting, and social orientation. The median mental age is obtained for each section of tests. The median of the nine sections gives the mental age of the child. The Intelligence Quotient is computed from this.

Before constructing his scale, Hiskey made a preliminary study of the deaf in school; then a survey of the

---

[10] Poull, Louise E., The Randall's Island Performance Series, Columbia University Press, 1931, p. 1.

[11] Schick, Helen, "The Use of a Standardized Performance Test for Pre-School Age Children With a Language Handicap", Proceedings of the International Congress on the Education of the Deaf, West Trenton, N.J., July, 1940.

literature pertaining to the deaf, and a study of almost every kind of intelligence or learning aptitude test item that has been used. Preference was given to the item that met certain criteria:

1. Was it the type of item which experience had shown to yield high correlation with acceptable criteria of intelligence or learning ability?
2. Was it the type of item which could be included in a non-verbal test?
3. Was the item similar to the tasks, or task, which the young deaf child did in school?
4. Could the item be presented in such a way that directions could be given through simple pantomime?
5. Could the item be constructed and presented in such a way that the child could make a definite response, thus making the scoring objective and easily done?
6. Would the item be appealing or attractive to the subject?
7. Could the item be scored without the score being based on time?
8. Did the difficulty of the item appear to be within the age range of the standardizing group?
9. Did the item seem likely to show a high discriminative capacity?[12]

Selection was made on the basis of the first criterion. Terman and Merrill have pointed out that,

> Prominent among tests which have universally proved their worth are analogies, opposites, comprehension, vocabulary, similarities and differences, verbal and pictorial completion, absurdities, drawing designs from copy and from memory, memory for meaningful material and for digits.[13]

---

[12]Hiskey, Marshall S., A Non-Verbal Test of Learning Aptitude Especially Adapted for Young Deaf Children, A Dissertation, Lincoln, Nebraska, July, 1940, p. 10.

[13]Ibid., p. 10-11.

Items of speed or the timed type were not used. Items of the same type were grouped and arranged in order of difficulty so that each group became a small power test in itself. All the items were of a type which had been used before but few of them had ever been assembled in the manner which they appear in the Hiskey scale.[14]

The age norm has been used in the Hiskey scale as the means of interpreting test results, but the term "Mental Age" has not.

> Although the high correlation which exists between intelligence and learning ability would make the use of the term a justifiable one, the fact that numerous items of the test have been adopted because of their similarity to the abilities which the deaf child must exhibit in school, and because the use of Mental Age would undoubtedly tend to suggest a Binet Mental Age which in turn would suggest the corresponding M.A. of the hearing child and thus lead to false comparisons, the term "Learning Age" is used instead.[15]

A learning age of five just means that according to the results of the test the child is able to do those things which the average <u>deaf</u> child of five years is able to do.

Hiskey states that there are reasons to question the use of the "Learning Quotient" as an index of ability of the deaf child, but there are those who wish to use it. He cautions the use until we have more reliable information regarding the constancy of the measure. The Learning Quotient is derived by dividing the Learning Age by the Chronological

---

[14]Ibid., Chapter II.

[15]Ibid., p. 20.

Age and is comparable to the Intelligence Quotient.[16]

Test 1. Memory for Colored Objects

The material consists of two sets of eight colored sticks each; one set placed in front of the child and the other set for the examiner. For practice purposes, and to insure that the subject understands that he is to select sticks of the same color as those which the examiner puts forward, a stick is placed a few inches in front of the subject's row of sticks. The examiner indicates that the subject is to put the stick of the same color near the former's stick. Go through the set of sticks this way, but do not score the responses. Then place a cardboard screen over the subject's sticks, put forward a stick in the order indicated in a list, and attract the subject's attention to it. Remove the screen and place it over the examiner's stick and indicate the subject is to select his stick of the same color. There are eight series of one stick each, three series of two sticks, two series of three, two series of four, and two series of five colors. The score is the total number of parts correctly reproduced from memory. If the subject fails two complete series, the examiner need not go on.

Test 2. Bead Stringing.

A box of thirty-six bright red beads (round, oblong, and cylindrical) and two strings are used. This test

---

[16] Ibid., Appendix on Learning Quotient.

is divided into two sections. The first part consists of counting the number of beads the subject can put on his string in two minutes. This part does not enter into the scoring if the subject is successful in making one of the patterns in the second part of the test. The subject makes pattern one and two with the examiner's pattern before him; pattern three, four, and five are made from memory. The score here is the total number of patterns correctly reproduced.

Test 3. Pictorial Associations

This test consists of twelve series of mounted pictures. In each series two pictures are mounted side by side and a recess is left for the insertion of a third picture. There are four individual pictures with each series of the mounted pictures, and the subject is to fill the recess with the correct one. The first series, for example, shows a boy riding on a scooter, a boy riding on a tricycle and the four individual pictures of a monkey, a boy riding in a wagon, a flower, and a fish from which the subject is to make the correct selection to put in the recess. He is to associate the boy riding in a wagon with the boys riding on a scooter and on a tricycle. The possible score is twelve.

Test 4. Block Building

There are sixteen yellow blocks, and eight cards showing block patterns for this test. The patterns are presented one at a time to the child and the examiner indicates that he should make them with his blocks. The examiner is permitted to build the first two patterns for the subject and let him copy that pattern. If he misses three consecutive designs, it is not necessary to give the remaining ones. The score is the total number of designs completed. The designs are scored as correct if the subject reproduces the pattern with a reasonable degree of alignment.

Test 5. Memory for Digits

The material consists of two sets of arabic digits, one through nine. Numbers one to five are placed before the subject, the first practice number (4) is placed before him, and he is signalled to place his number four beside it. Go through the practice numbers 4, 1, 5, 2, 3, and then through the practice numbers 51, 42, 35. The series of two numbers must be placed in the same order as those of the examiner. Then the cardboard screen is placed over the subject's numbers, the examiner places his first number forward, covers it, and indicates the subject is to place the corresponding number to the side. All five of the individual numbers are given, then three series of two numbers each. One success out of the three trials is all that is necessary, but correct selections of numbers are recorded whether they are arranged in order or not. Before presenting the three series of three numbers each, the examiner places the numbers 6, 7, 8, and 9 before the subject. The following three series are of four numbers each, and the last series of three are of five numbers each.

For the five series of one number the score is the number of digits correctly selected. For the series of 2, 3, 4, and 5 numbers each, where the subject has reproduced correctly any one series, he is credited with two points. On those parts where the subject selects the right digits in one or more series but is not able to arrange them in the proper order, the score is one for the first, such part, but full credit of two points for all such subsequent parts. Only once is the subject penalized for not getting the proper arrangement of the digits.

Test 6. Completion of Drawings

A sheet of paper containing fifteen pictures, each with a part missing, is placed before the subject. Picture one (a chair with a leg missing) is a practice exercise and the examiner is permitted to help the subject draw in the missing leg. The subject is to proceed through the rest of the test alone. The score is the number of drawings (other than number one) correctly completed.

Test 7. Pictorial Identification

Six sets of five pictures each, mounted on cardboard, and twenty-four individual pictures for matching purposes are provided for this test. Set one is placed before the subject, and then the four individual matching pictures one at a time. The subject is to place each one of the individual pictures on the correct picture in the set. Each single picture is removed after the subject makes his response. The score is the total number of correct selections.

Test 8. Paper Folding

The examiner keeps one six-inch square of paper and gives another one to the subject. He folds his paper according to pattern one and the subject is to fold his in the same manner. There are seven patterns, and the score is the number of patterns correctly folded.

Test 9. Visual Attention Span

Mounted pictures are used for this test: three singles, two doubles, one series of three, one series of four, one series of five, and one series of six; and fifteen individual pictures. The fifteen individual pictures are arranged before the subject in three rows of five each according to the numbers on the backs of the pictures, and the subject is given an opportunity to look at them. Then the examiner covers them with the cardboard screen, and shows the subject the mounted single picture number one. The picture is then withdrawn and the screen removed, and the subject is to find the correct picture. All three single pictures are presented and then the group containing two pictures. For this and the remaining groups point to each picture consecutively three times. The total possible score is six points. Thus if a subject gets only one of the three single pictures, he gets credit for that series. The same is true for the two sets of two pictures each. The remaining groups have only one series each, thus demanding perfect selections on each to obtain credit.

Test 10. Puzzle Blocks

The material consists of one solid, colored block (two opposite sides of red, two blue, and two of yellow) and eight colored blocks that have been divided into parts. The solid block is the model and remains in the center of the table throughout the test. The examiner puts puzzle number one in front of the subject, and indicates that he is to put it together. The examiner can help because it is just a practice exercise and is not counted in the score. Block number two is then presented and when the subject has put it together, it is pushed over beside number one and permitted to remain on the table. The same procedure is followed with all the completed blocks. The recommended time limits are two minutes for each of the first three blocks and four minutes for each of the last four blocks. These are very liberal time limits and are only recommended so that the slow subject will not unduly prolong the test. If the subject has not completed the block at the end of the time limit, the examiner may complete it for him and push it over with the other completed blocks. After he has failed three consecutive blocks, there is no need to present additional ones. The score is the number of blocks completed regardless of color. The color is chiefly for making the blocks attractive and to give added cues as to where the parts belong.

Test 11. Pictorial Analogies

There are ten series of mounted pictures. In each set three pictures are mounted; two of the pictures are grouped on one side of a division strip and the third picture is mounted on the other side of the division where there is also a recess for the insertion of a fourth picture. There are four individual pictures for each series. The examiner holds the mounted pictures in the left hand and points to the first picture with the first finger of the right hand then to the second picture. Then using the first and second fingers he points to the two pictures simultaneously. Repeat this with the third picture and the recess. This is to help the child understand that the two parts go together. The subject is to put one of the individual pictures in the recess beside the third picture. The procedure is the same for the other series. The possible score is ten. One of these series, for example, contains pictures of a girl and a woman on one side of the division strip, and a boy and the recess on the other side. The possible choices are a girl in a swing, a man, a shoe, and a boy on a sled. The correct analogy is the girl and woman, and the boy and man.

At the completion of each test the examiner is to simply encircle the score made by the subject on the prepared score sheet. When all the tests have been completed, find the child's median learning age.

Some of the tests do not function over the entire age range and therefore, it is not always necessary to give them. Hiskey adopted the following rules:

> (1) Unless the subject's performance on the various items equals the norms for five-year-olds (or above) on at least half of the items, do not give him the "completion of drawings", "puzzle blocks", or "pictorial analogies" items. Thus he must equal the five-year norms on at least four of the remaining eight items before it is advisable to present the above three items. With the younger group (six and below) these three items should be administered last. (2) Unless the subject has three or more scores whose L.A. is below eight, do not give the "paper folding" and "picture identification" items. With the older group the remaining nine items would be given first. In most instances those individuals between the ages of six and eight years can be given all items of the scale.[17]

Procedure

The Nebraska Test of Learning Ability was given to sixty-one deaf children, ranging in age from four years one month to ten years five months at the time of the test, except for one retarded child whose chronological age was eleven years two months. Twenty-four of these deaf children were enrolled at St. Joseph's Institute for the Deaf in St. Louis County, Missouri, and the remainder attended Central Institute for the Deaf in St. Louis, Missouri. Each of the deaf children was also given either the Advanced Performance Scale or the Randall's Island Performance Series.

The children from St. Joseph's were taken to a class room not in use in the school and after the establishment of rapport between the examiner and the subject, the learning

---

[17] Hiskey, op. cit. Handbook, p. 10.

test was administered. In no case was there any behavior problem encountered. All of the children seemed quite eager to play the games with the examiner, although they were not accustomed to a testing situation. In each case the learning test was administered first and then after a time interval ranging from two days to one month eleven days the same children were tested on a performance scale. The same room was used.

The deaf children at Central Institute were taken by the same examiner to a room set aside for psychometric testing. All were acquainted with the examiner and the majority of the children were familiar with the testing situation. In all these cases the performance scale preceded the learning test. Since each child is given a mental test each year at Central Institute, there were some children with as many as six test scores. In a case where a child had only one score, this score was listed as the Intelligence Quotient for the child. For a child with more than one score, the average of all test scores was found and the average Intelligence Quotient was listed. The time interval between the learning test and the most recent performance test ranged from one day to four months twenty-five days.

Sixty-six hearing children were taken either from the Speech Correction Department at Central Institution or from the speech clinic for children enrolled in the regular public schools. The age range at the time of the Hiskey test was four years five months to ten years seven months, except for one very bright child whose chronological age was three years three months, and three older retarded children. In most cases the performance test preceded the learning test. Again, the average Intelligence Quotient was used for those children having two or more performance test scores. For those in the Speech Correction Department the time interval between the learning test and the most recent performance test ranged from one day to nine months twelve days, an average of two months; the range for speech clinic children was from four days to sixteen months twenty-four days, an average of three months seven days.

The speech patterns of the hearing children were not normal. They were characterized by such speech defects as: oral inaccuracies, stammering, stuttering, delayed speech, spastic speech, and motor aphasia. Sensory aphasics were not included because of their lack of understanding.

The mean, median, and range of Intelligence Quotients for the performance test and the mean, median, and range of Learning Quotients for the learning test were found for the deaf and hearing groups seperately. The standard deviations of the Intelligence Quotients and of the Learning Quotients were also computed.

The correlation coefficients were computed between test scores on the Nebraska Test of Learning Ability and on the Performance Scales for both the deaf and hearing groups to determine the relationship of the two types of tests. The probable error of the correlation coefficients was also computed.

Twenty-six children enrolled at Central Institute (of whom seven were hearing) were retested on the Nebraska Test of Learning Ability, the time interval ranging from eight and one-half months to thirteen months. The mean and median Learning Quotient and the standard deviation of the Learning Quotient were computed for both the first test and the retest and comparisons were made. The correlation coefficient between the test scores of the first test and the retest, and the probable error of the coefficient were also determined.

The gain in months in Learning Age was calculated for each child retested on each of the individual tests, and from these figures the mean gain in each test was computed. The individual test scores reaching a maximum on both the first test and the retest were eliminated in computing the mean gain.

The Hiskey test was then studied to determine the increasing difficulty at each age level for the hearing and deaf groups combined.

To find which tests were the most difficult for the deaf children and which were most difficult for the hearing children, Subject Quotients were computed for each child for each individual test, and then the mean Subject Quotient for each individual test was computed. The tests were listed in order of difficulty separately for the deaf and the hearing for purposes of comparison.

Finally, the number of points variation between the scores of each child on the Nebraska Test of Learning Ability and the Performance Scales were tabulated.

Chapter IV

RESULTS

The Nebraska Test of Learning Ability was constructed by Hiskey to measure the learning ability of deaf children. Prior to this study the test had not been applied to hearing children. No comparison had been made between the learning quotient, as obtained on this test, and the intelligence quotient. The test was given to hearing children to determine their scores on a test designed for the deaf.

Table I

Comparison of the Learning Test with the Performance Scale

	N	r	$P.E._r$
Deaf	61	.74	.04
Hearing	66	.90	.02
Total Group	127	.86	.02

Table I shows a comparison of the Learning Test with the Performance Scale. The correlation coefficient between scores on the two tests for sixty-one deaf children was .74 $\pm$ .04, for the hearing children the coefficient was .90 $\pm$ .02, while for the total group it was .86 $\pm$ .02. The probable errors are very low, indicating high statistical reliability for the test. The high correlation coefficients show that the two tests are measuring the same thing; the Learning Test seems to be another type of intelligence test.

Table II

Comparison of Variation between the Learning Test and the Performance Scale

Variation	Percent of Deaf	Percent of Hearing	Percent of Combined Group
Less than 5 points	27.8	37.8	33.1
Less than 10 points	54.1	54.5	54.3
10-19 points	21.3	33.3	27.6
Over 20 points	24.6	12.1	18.1

Table II indicates variation between scores on the Learning Test and the Performance Scale in terms of points. Fifty-four and three tenths percent of the Learning Test scores varied less than ten points from the scores on the performance test. This furnishes additional evidence that the Learning Test measures the same type of ability as the performance scale. Twice as many deaf as hearing children showed a variation of over twenty points.

Table III

Comparison of Deaf and Hearing on the Learning Test

	N	Range of C.A.	Mean L.Q.	Median L.Q.	Range of L.Q.
Deaf	61	4-1 -- 11-2	113.87	113	68-161
Hearing	66	3-3 -- 13-7	101.67	105.5	43-167

Table IV

Comparison of Deaf and Hearing on the Performance Scale

	N	Range of C.A.	Mean I.Q.	Median I.Q.	Range of I.Q.
Deaf	61	3-8 -- 10-10	116.62	117	71-168
Hearing	66	3-1 -- 13-10	101.05	99.5	46-169

Table III shows a comparison of the deaf and hearing groups on the Learning Test. The mean Learning Quotient of 113.87 shows the deaf to be a decidedly superior group. There was no selection of cases made in the present study that would have influenced the mean Learning Quotient. All the deaf children between the chronological age range of four to ten were used, with the exception of one retarded child whose chronological age was eleven years two months.

The hearing group seems to be an average group in so far as mean Learning Quotient is concerned. Their range is wider because of the inclusion of scores of some speech defectives whose lack of speech is caused by a mental retardation. Several spastics have also been included in the study, but their scores would not have influenced the mean since the Learning Test is not a timed test. Again, all children within the desired chronological age range were used, except for one unusually bright child who was three years and three months old, and several older retarded children.

Since there was no matching of the deaf and hearing groups, it is not possible to make any statement concerning the relative learning ability of the two groups.

Table IV shows the comparison of the deaf and hearing on the Performance Scale. The majority of the children were tested on the Advanced Performance Scale, but a few were tested on the Randall's Island Series. White found a correlation coefficient between these two scales of .65 ± .04.[1] The mean Intelligence Quotient of the deaf group shows them to be above normal, while the mean Intelligence Quotient of the hearing group is nearly the same as the mean Learning Quotient. The range of the Intelligence Quotients is larger for the hearing than for the deaf, and is due to the inclusion of scores of some mentally retarded children.

Figure one shows a comparison of the frequency distribution of deaf and hearing Learning Quotients. The Learning Quotients are plotted in step intervals of ten along the abscissae, and the frequency of occurrence of these quotients on the ordinates. The negative skewness of the curve for the hearing shows a slight tendency for the scores to be massed at the high end of the scale, while the positive skewness of the scores for the deaf shows a very slight tendency for the scores to be massed at the low end of the scale.

Figure two is plotted in the same manner as Figure one, but shows a comparison of the distribution of Intelligence Quotients of the deaf and hearing. The curve of Intelligence Quotients of the hearing shows a positive skewness; the curve for the deaf a negative skewness.

The Learning Quotients of both the deaf and hearing are combined and plotted in Figure three; the curve of the

---

[1]White, Jennylouise, A Performance Test for School Age Children With a Language Handicap, M.S. Thesis, Washington University, 1939, p. 32-47.

combined Intelligence Quotients is superimposed on the curve of the Learning Quotients. The two curves show a normal distribution and an equal range. The curve of Intelligence Quotients is peaked at the step interval of 110-119, while the greatest number of Learning Quotient scores occurred at the interval of 120-129.

Table V

Comparison of the First Test
and the Retest on the Learning Scale

	N	Mean L.Q.	Median L.Q.	Standard Deviation
First test	26	114.42	116.5	25.93
Retest	26	113.96	116	24.34
$r = .87 \pm .03$				

Twenty-six children were retested on the Learning Scale, and Table V shows a comparison of the results. A correlation coefficient of $.87 \pm .03$ was computed between the two tests, indicating high reliability. The means, medians, and the standard deviations are very close for the test and retest.

Figure I

Comparison of Frequency Distribution of Deaf and Hearing Learning Quotients on Hiskey Test

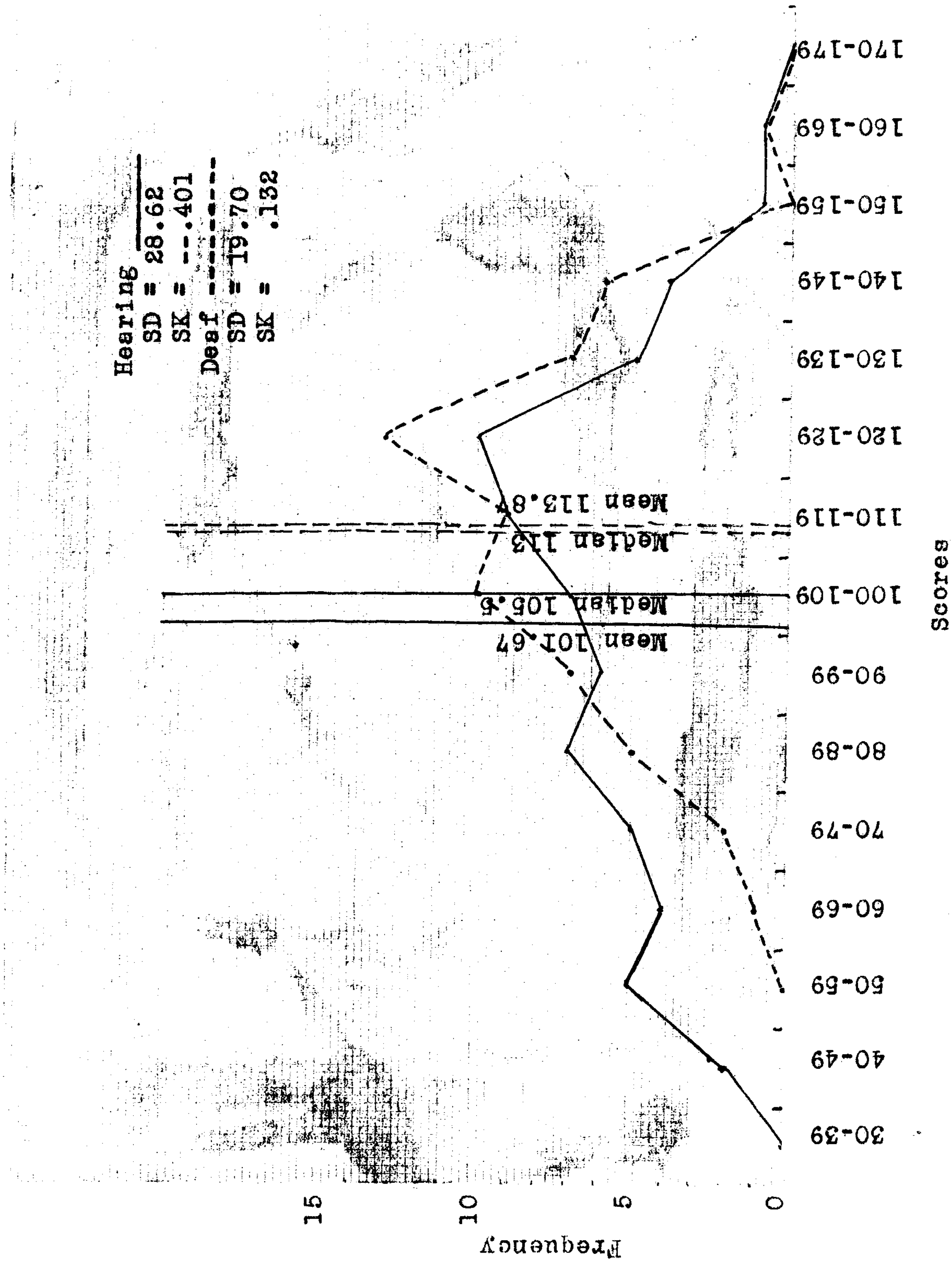

Figure II

Comparison of Frequency Distribution of Deaf and Hearing Intelligence Quotients on Performance Scales

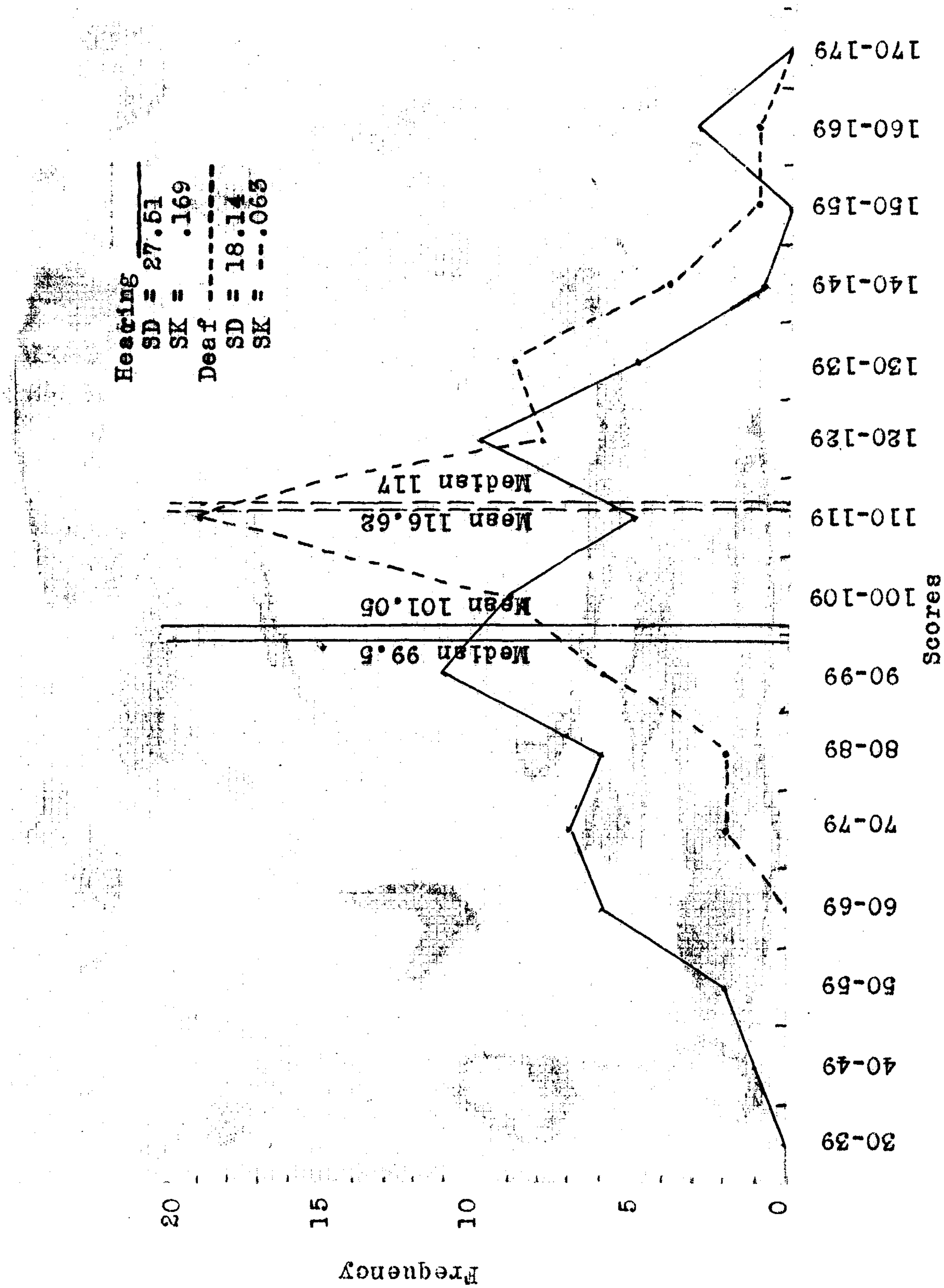

Figure III

Comparison of Frequency Distribution of Deaf and Hearing Learning Quotients and Deaf and Hearing Intelligence Quotients

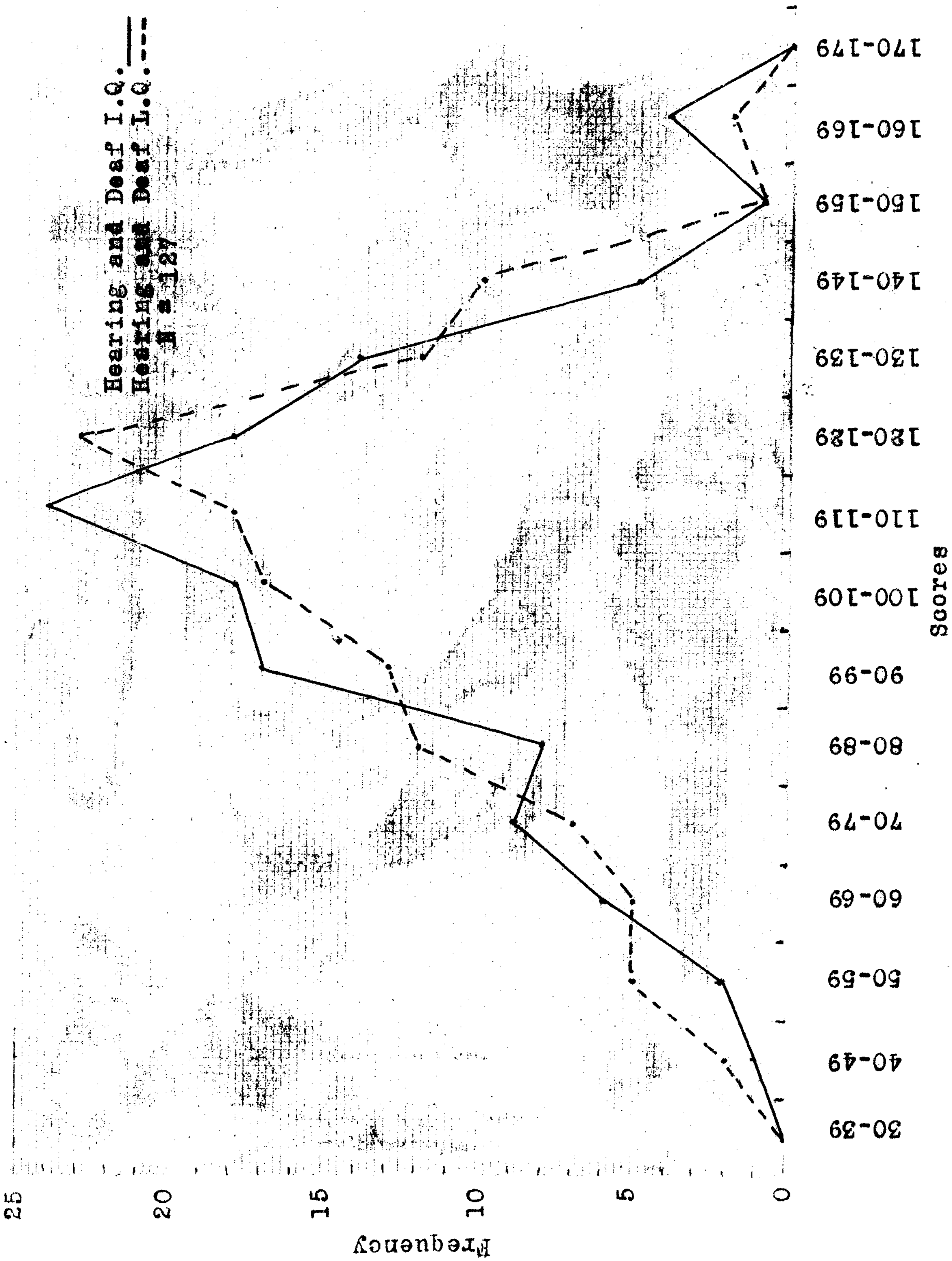

Table VI

Improvement in Months of
Learning Age Between First Test and Retest

Name of Test	Mean Improvement in Months of L.A.
1. Memory for colored objects	15.5
2. Bead stringing	4.9
3. Pictorial Association	15.8
4. Block Patterns	5.3
5. Memory for Digits	15.0
6. Completion of Drawings	14.1
7. Pictorial Identification	6.9
8. Paper Folding	- 3.8
9. Visual Attention Span	12.0
10. Puzzle Blocks	9.7
11. Pictorial Analogies	4.6
Battery of Tests	9.1
Mean Increase in C.A. between tests	10.8

Table VI shows the improvement in months of learning age between the first test and the retest on the Hiskey test. The mean increase in chronological age between tests is 10.8. Memory for Colored Objects, Pictorial Association, Memory for Digits, Completion of Drawings, and Visual Attention Span show an average increase greater than the mean increase in chronological age. They show improvement above normal expectancy. The average increase in months of learning age is below the mean increase in chronological age for Bead

Stringing, Block Patterns, Pictorial Identification, Paper Folding, Puzzle Blocks, and Pictorial Analogies. The mean improvement for all the tests is 9.1, but if the Paper Folding score is omitted the total mean improvement is 10.4 which compares favorably with the mean increase of 10.8 in chronological age.

Paper Folding is the only test which shows no mean improvement; instead it shows a loss. The author believes this negative score may be due to the administration of the test itself. It is difficult to make the subject understand that he is to watch the examiner until he has finished folding the pattern and then fold his own. There is confusion when the examiner attempts to stop the subject from folding his paper, and this results in loss of memory for the pattern. Even putting the paper out of the subject's reach does not alter the situation because in his eagerness to start folding his paper he will try to reach for the paper. Then too, the number of cases to whom this test was administered is small. Paper Folding is administered only to those children having three scores below a learning age of eight. Only eleven of the twenty-six children retested on the Learning Test were tested on Paper Folding.

In order to determine the increasing difficulty of the Learning Test for each six months period, the deaf and hearing were divided into groups according to Chronological Age; those whose Chronological Age was from four years to four years five months in the first group; those whose Chronological Age was from four years six months to four years eleven months in the next group, and so on up through the Chronological Age group of ten years to ten years five months. The mean Chronological Age, mean Learning Age, and the mean Learning Quotient was calculated for each group, and then the gain in months of Chronological Age, Learning Age, and points of Learning Quotient was found between the first and second group, between the second and third group, between the third and fourth group, and on up between the groups.

Table VII

Increasing Difficulty of the Test for Each Six Month Interval

Mean Chronological Age	Mean Learning Age	Mean Learning Quotient	N	Gain in Months Chronological Age	Gain in Months Learning Age	Gain in Points Learning Quotient
4-3	5-3	116	9			
4-9	5-10	117	8	6	7	1
5-3	6-5	129	12	6	7	12
5-9	5-7	93	8	6	-10	-36
6-3	6-7	102	8	6	12	9
6-8	8-2	109	8	5	19	7
7-3	8-2	113	10	7	0	4
7-10	8-5	106	8	7	3	- 7
8-2	7-11	108	10	4	- 6	2
8-9	10-0	118	13	7	25	10
9-3	8-5	92	11	6	-19	-26
9-10	9-3	104	11	7	10	12
10-2	9-11	109	5	4	8	5

Table VII shows the increasing difficulty of the Learning Test for each six months interval. Only 121 of the 127 cases were used. The three year old child's score is not included and the scores of those ten years six months and older are not included. Between the mean chronological age of four years three months and four years nine months there is a gain of seven months in learning age compared with a gain of six months in chronological age. There is the same amount of gain between the mean chronological age of four years nine months and five years three months. Between the mean chronological age of five years three months and five years nine months there is a loss of ten months in learning age. The test appears to be too difficult for these children in the present study. Between five years nine months and six years three months, and then between six years three months and six years eight months there is a gain of twelve and nineteen months in learning age, indicating that the test is easy at these levels. From the mean chronological ages of six years eight months to seven years three months, from seven years three months to seven years ten months, and from seven years ten months to eight years two months there is no gain, with a slight gain of just three months and then a loss of six months in learning age respectively, indicating that the test is difficult at these three age groups. At eight years nine months there is a gain over three times as great as the gain in chronological age. The test seems easiest at this age level. It seems hardest at nine years three months where there is a loss of nineteen months in learning age, but this may not be a true loss when the gain of twenty-five months in the previous six month interval is taken into consideration. Again between nine years three months and nine years ten months, and from nine years ten months to ten years two months there is a gain of ten and eight months in learning age respectively, a little better than average gain.

The author can see no reason for the spotted gain and loss of months of learning age. Perhaps it can be attributed to one of the following factors or perhaps to both: (1) technical faults in the construction of the test, making it too easy at some levels and too difficult at others; or (2) differences in the educational opportunities of the group tested.

Table VIII

Difficulty of the Individual Tests of the Learning Test

Name of Test	Mean Subject Quotient		Rank in order of Difficulty	
	Deaf	Hearing	Deaf	Hearing
1. Memory for Colored Objects	113.26	105.09	6	5
2. Bead Stringing	104.67	96.14	11	9
3. Pictorial Association	109.13	98.45	8	7
4. Block Patterns	117.92	106.17	4	4
5. Memory for Digits	109.34	104.14	7	6
6. Completion of Drawings	119.75	116.78	3	2
7. Pictorial Identification	120.56	92.81	2	10
8. Paper Folding	106.80	80.33	10	11
9. Visual Attention Span	107.31	97.38	9	8
10. Puzzle Blocks	117.45	113.48	5	3
11. Pictorial Analogies	126.48	117.47	1	1

Table VIII shows the mean subject quotient for each of the individual tests of the Learning Test and their rank in order of difficulty for the deaf and the hearing. It is interesting to note that Pictorial Analogies ranks first with both the deaf and hearing groups; that is, it is the easiest for both groups. There are no more than two places difference in rank of the individual tests for the deaf and hearing except for Pictorial Identification which ranks second for the deaf and tenth for the hearing. The education of the deaf is more heavily weighted with picture identification and matching than the education of the hearing, and this may account for the difference in rank. The mean subject quotient of each individual test is above average for the deaf, Pictorial Analogies, Pictorial Identification, and Completion of Drawings ranking highest. Tests showing scores above average for the hearing are: Memory for Colored Objects, Block Patterns, Memory for Digits, Completion of Drawings, Puzzle Blocks, and Pictorial Analogies. Pictorial Analogies, Completion of Drawings, and Puzzle Blocks rank highest for the hearing. Paper Folding ranks far below the other tests for the hearing with a mean quotient of 80.33. Others ranking below one hundred are: Bead Stringing, Pictorial Association, Pictorial Identification, and Visual Attention Span.

Chapter V

SUMMARY AND CONCLUSIONS

1. Sixty-one deaf children were tested on the Nebraska Test of Learning Ability. Sixty-six speech defective children with normal hearing were also tested on the Learning Test. The above group was given either The Advanced Performance Scale used at Central Institute for the Deaf or the Randall's Island Series for younger children. Comparisons were made between the Learning Test and the Performance Scale of Intelligence to determine the relationship between the two types of tests.

2. The correlation coefficient between scores on the Learning Test and the Performance Scale was .74 $\pm$ .04 for the deaf and .90 $\pm$ .02 for the hearing, indicating that the Learning Test is a valid test of intelligence.

3. Fifty-four and three tenths percent of the Learning Test scores varied less than ten points from the scores on the Performance Test, furnishing additional evidence that the Learning Test measures the same type of ability as the Performance Scale.

4. The sixty-one deaf children showed a mean Learning Quotient of 113.87 and a median Learning Quotient of 113, with a range of sixty-eight to one hundred sixty-one. The range of chronological age was four years one month to eleven years two months. The sixty-six hearing children showed a mean Learning Quotient of 101.67 and a median Learning Quotient

105.5, with a range of forty-three to one hundred sixty-seven. The chronological age range was three years three months to thirteen years seven months.

5. The mean Intelligence Quotient of the deaf was 116.62, the median Intelligence Quotient 117, and the range seventy-one to one hundred sixty-eight. For the hearing the mean Intelligence Quotient was 101.05, the median 99.5, and the range forty-six to one hundred sixty-nine.

6. Twenty-six children were retested on the Learning Scale. The correlation coefficient of .87 $\pm$ .03 indicated high reliability. The mean Learning Quotient on the first test was 114.42, the median 116.5, and the standard deviation 25.93. For the retest the mean Learning Quotient was 113.96, the median 116, and the standard deviation 24.34.

7. The mean improvement in months of learning age for the individual tests of the Learning Scale was computed. The mean increase in chronological age between the first test and the retest was 10.8. Memory for Colored Objects, Pictorial Association, Memory for Digits, Completion of Drawings, and Visual Attention Span showed an average increase between tests greater than the mean increase in chronological age; Bead Stringing, Block Patterns, Pictorial Identification, Paper Folding, Puzzle Blocks, and Pictorial Analogies less than mean increase in chronological age. The mean improvement for all the tests was 9.1.

8. The increasing difficulty for each six months

interval was computed for the Learning Test. The gain in months of learning age from one interval to the next was not consistent. At some intervals the gain in months of learning age was far greater than the gain in months of chronological age, and at some intervals the gain was far less than the gain in months of chronological age.

9. The mean subject quotient of each individual test of the Learning Test was above average for the deaf, Pictorial Analogies, Pictorial Identification, and Completion of drawings ranking highest. The mean subject quotients above average for the hearing were: Memory for Colored Objects, Block Patterns, Memory for Digits, Completion of Drawings, Puzzle Blocks, and Pictorial Analogies. Pictorial Analogies, Completion of Drawings, and Puzzle Blocks ranked highest in subject quotients for the hearing.

The results as a whole indicate normalcy in learning ability and in intelligence for both the deaf and hearing. The deaf are not as retarded as studies of others would indicate.

The writer feels that the Nebraska Test of Learning Ability is heavily weighted with visual memory. Five of the eleven tests involve this mental faculty: Memory for Colored Objects, memory for bead patterns in the Bead Stringing test, Memory for Digits, memory for patterns in the Paper Folding Test, and the Visual Attention Span test. Since the hearing handicap of the deaf entails an emphasis upon visual education and visual aids to learning, it is possible that the visual weighting of the Learning Test has given the deaf an advantage over the hearing. This advantage may be the reason for the higher mean Learning Quotient of the deaf. On the other hand, tests designed for the hearing are weighted in language and place the hearing at an advantage. It would be desirable also to have hearing norms on the Learning Test in order to determine the true effect of this visual weighting. Until the qualities tested by the various items have been more clearly defined by factorial analysis further discussion is unjustified.

Another objection is the length of time required for administration of the test. The average length of time spent in administering the test to the 127 subjects was one hour and ten minutes. Only twenty-one of the subjects finished before one hour had elapsed. Three children required forty-five minutes to complete the test, and on the other hand, three required one hour and thirty-five minutes. One very slow child needed one hour and forty minutes. The latter tests, of course, were divided into two sittings. In spite of the length of the test, interest was maintained even for the younger children. All the children seemed to enjoy "the games".

The chief advantages of the scale are: (1) the elimination of so many speed tests; (2) the complete absence of oral instructions and responses; (3) the ease of administration and scoring; (4) the attractiveness of the tests; (5) the insight into the educational background of the child since items in several instances have been selected because of the close relationship to tasks actually done in school; (6) the use of the test with young deaf children; and (7) the portability of the material.

It would be of value in further study to give the Learning Test to hearing children without speech defects and correlate these scores with the scores of the deaf and with scores of the hearing with speech defects; also to compare the scores on a language test with scores on the Learning Test.

The results of this study further substantiate the opinion of other investigators that a testing program should be instituted in every school for the deaf with annual retests for each child. It is desirable to administer both an intelligence test and the Nebraska Test of Learning Ability in order to give the teacher and administrator as broad a measure of the child as possible.

BIBLIOGRAPHY

Baker, Harry, and Leland, Bernice, Detroit Tests of Learning Aptitude A New Instrument of Mental Diagnosis, Bloomington, Ill: Public School Publishing Co., 1935.

Bayley, Nancy, "Mental and Motor Development From Two to Twelve Years", Review of Educational Research 9: 18-37, 1939.

Bishop, Helen M., "Performance Scale Tests Applied to Deaf and Hard of Hearing Children", Volta Review 38:447, 1936.

Blatz, W. E., "Beyond the I.Q.: Some Comments on Intelligence", Education Digest, October 1939, p. 24-26.

Boynton, Paul, Intelligence Its Manifestations and Measurement, New York: D. Appleton and Co., 1933, p. 1-30, 149-206, 426-441.

Brill, Moshe, "Performance Tests as Aids in the Diagnosis of Maladjustment", Pedagogical Seminary 49:199-214, 1936.

Brill, Richard, "Latest Developments in Testing in Schools for the Deaf", Report of the Proceedings of the Thirty-Second Meeting of the Convention of American Instructors of the Deaf, Fulton, Missouri, June 1941, p. 393-396.

Buros, Oscar, Mental Measurements Yearbook 1940. New Brunswick: Rutgers University Press.

Cattell, Psyche, "Intelligence Tests", Review of Educational Research 8:221-228, 1938.

Cutler, Elizabeth, "Summary of Psychological Experiments with the Deaf", American Annals of the Deaf 86:181-192, March, 1941.

Drever, James, and Collins, Mary, Performance Tests of Intelligence A Series of Non-linguistic Tests for Deaf and Normal Children, Edinburgh: Oliver and Boyd, 1928.

Goodenough, Florence, "Can We Influence Mental Growth: A Critique of Recent Experiments", Educational Record 21:120-143, 1940.

Hiskey, Marshall S., A Non-Verbal Test of Learning Aptitude Especially Adapted for Young Deaf Children, Unpublished Dissertation, University of Nebraska, Lincoln, Nebraska, July, 1940.

Husband, R. W., "The Influence of a Wide Range in Intelligence on Inter-correlations Among Learning Abilities", Abstract, Psychological Bulletin 37:580, 1940.

__________, "Intercorrelations Among Learning Abilities: III. The Effect of Length of Tests Upon Intercorrelations", Journal of Genetic Psychology 58:427-430, 1941.

Johnson, Helmer, "I.Q. Skullduggery?" School and Society 57:269-272, 1943.

Jones, H. E., "Relationships in Mental Development", National Education Association Proc., 1939, p. 90-91.

Keys, Noel, "Applications of Intelligence Testing", Review of Educational Research 8:241-256, 327-334, 1938.

Lane, Helen Schick, and Schneider, Jennylouise Lockwood, "A Performance Test for School Age Deaf Children", American Annals of the Deaf, 86:441-447, November, 1941.

Maxfield, Francis, "Substitutes for the 'I.Q.'", Educational Research Bulletin, 15:92-94, 1936.

__________, "Trends in Testing Intelligence", Educational Research Bulletin 15:134-141, 1936.

Morris, Charles, "A Critical Analysis of Certain Performance Tests", Journal of Genetic Psychology 54:85-105, 1939.

Morrison, W. J., "The Ontario School Ability Examination", American Annals of the Deaf, p. 184-189, March 1940.

Newlee, Clara, "A Report of Learning Tests With Deaf Children", Volta Review 21:216-223, 1919.

Peterson, Edwin, and Williams, James, "Intelligence of Deaf Children As Measured by Drawings", American Annals of the Deaf 75:273-290, 1930.

Pintner, Rudolf, Intelligence Testing, New York: Henry Holt and Co., 1923, p. 1-88, 93-135, 312-327.

__________, "Contributions of Psychological Testing to the Problems of the Deaf", Proceedings of the International Congress on the Education of the Deaf, West Trenton, N.J., June 1933, p. 213-220.

Pintner, Rudolf, Jon Eisenson, and Stanton, Mildred, The Psychology of the Physically Handicapped, New York: F. S. Crofts and Co., 1941, p. 101-130.

Pintner, Rudolf, and Paterson, Donald, "The Survey of a Day School for the Deaf", American Annals of the Deaf 61:417-433, Nov., 1916.

__________, __________, "A Class Test With Deaf Children", American Annals of the Deaf 61:264-275, 1916.

__________, __________, "The Form Board Ability of Young Deaf and Hearing Children", American Annals of the Deaf 16:184-189, 1916.

__________, __________, "Psychological Tests of Deaf Children", Volta Review 19:661-667, 1917.

__________, __________, "A Comparison of Deaf and Hearing Children in Visual Memory for Digits", Journal of Experimental Psychology 2:76-88, 1917.

__________, __________, "Some Conclusions From Psychological Tests of the Deaf", Volta Review 20:10-14, 1918.

__________, __________, A Scale of Performance Tests, New York: D. Appleton and Co., 1931, p. 1-25.

Porteus, S. D., "The Measurement of Intelligence: Six Hundred and Fifty-Three Children Examined by the Binet and Porteus Tests", Journal of Educational Psychology 9:13-31, 1918.

Psychological Tests and Their Uses", Review of Educational Research 14:1-37 Feb. 1944.

Rockwell, John G., "Intelligence Testing: Its Basic Assumptions and Unanswered Questions", Educational Method 19:80-92, November, 1939.

Ruch, G. M., and Orata, P. T., "Testing: Intelligence, Aptitude, Personality, and Achievement", Review of Educational Research 9:514-523, 1939.

Schick, Helen, "The Use of a Standardized Performance Test for Pre-School Age Children With a Language Handicap", Proceedings of the International Congress on the Education of the Deaf, West Trenton, N. J., June 1933, p. 526-532.

__________, "A Performance Test for Deaf Children of School Age", Volta Review 36:657-659, 1934.

Schick, Helen F., and Meyer, Max F., "The Use of the Lectometer in the Testing of the Hearing and the Deaf", American Annals of the Deaf, 77, 4:292-304, Sept. 1932.

Schieffelin, Barbara, and Schwesinger, Gladys, Mental Tests and Heredity Including A Survey of Non-Verbal Tests, New York: Galton Pub. Co., 1930, p. 1-70, 86-106.

Stoddard, George, "The I.Q.: Its Ups and Downs", Educational Record 20:44-57, 1939.

______________, "The Intellectual Development of the Child", Educational Digest 5:17-20, April 1940.

Symonds, Percival, "Review and Preview", Review of Educational Research 8:217-220, 1938.

Terman, L. M., and Merrill, M. A., Measuring Intelligence. Chicago: Houghton Mifflin Co., 1937.

Thurstone, L. L., "A New Conception of Intelligence," Educational Record 17:441-450, 1936.

Tiege, Ernest W., "Breaking Down the I.Q.", Progressive Education, 13:603-605, 1936.

______________, Tests and Measurements in the Improvement of Learning. Chicago: Houghton Mifflin Co., 1939, p. 29-45.

White, Jennylouise, A Performance Test for School Age Children With a Language Handicap, M.S. Thesis, Washington University, St. Louis, Mo., 1939 (Also see Lane and Schneider)

Zeckel, A, and Van Der Kolk, J. J., "A Comparative Intelligence Test of Groups of Children Born Deaf and of Good Hearing, by Means of the Porteus Test", American Annals of the Deaf, 84:114-124, 1939.